INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2015, VOL. 61, NO. 4, PP. 403-408
Manuscript received March 10, 2015; revised November, 2015. DOI: 10.2478/eletel-2015-0053

# Connectivity Solutions in Automated Trading

Paweł Popławski

***Abstract*—The study analyzes the architecture and deployment of direct market access (DMA) solutions for automated trading of securities. It provides an overview of automated trading systems including: trading floor architecture, trading environment connectivity, and DMA solutions. Among a range of factors influencing operational capacities, round-trip latency has been recognized as the key quality differentiator of an automated trading floor. The study identifies potential opportunity costs due to latency levels as a major driver of technological progress in trading in highly liquid market conditions.**

***Keywords*—telecommunications, digital systems, systems architecture, automated trading, high-frequency trading**

## I. Introduction

AUTOMATED trading (AT) refers to transactions of buying or selling securities without necessary human participation, notably in the decision-making process. Based on Teall [1] an AT strategy is created when a trader or programmer designs a trading system for automated submission and allocation of trade orders among markets and over time so as to achieve an optimal price level. Owing to scalability and efficiency, the operational approach and technological requirements in AT and especially in high-frequency trading (HFT) significantly diverge from human-based trading.

Numerous research [2, 3, 4, 5] show that HFT orders tend to better reflect market information than traditional ones; furthermore, Admati and Peiderer [6] and Grossman and Stiglitz [7] had established that connection to direct information feeds facilitates to gain competitive advantage over other agents. It also enables traders for statistical arbitrage, defined by Lo [8: 260] as "highly technical short-term mean-reversion strategies involving large numbers of securities (hundreds to thousands, depending on the amount of risk capital), [and] very short holding periods (measured in days to seconds) [...]."

The architecture of a trading floor is determined by various factors. Firstly, it needs to be compatible with the technologies and software used in the organized securities markets (OSM) involved. Another determinant is the kind of DMA the trader has to each of the OSMs. Finally, entire trading environment has to comply with regulatory requirements.

The internal criteria relevant for execution quality control may require e.g. handling simultaneous price updates from several OSMs at specific rates, visibility into data freshness, and the presentation of evidence that the best-possible execution has been obtained. Finally, data traffic management has to optimize capacity utilization along with processing and default transit latencies. Latency monitoring thus needs to be carried out in near-real time, with sub-microsecond granularity of measurements, the ability to handle high message rates, and must differentiate application processing latency from network transit latency.

Institutional investors and brokers managing client accounts tend to show different attitudes towards trading [9, 10, 11] which suggests significant discrepancies regarding AT using DMA. Nonetheless, technology management represents a common challenge for executives in the financial industry. Issues range from inflexible IT budgets and depreciation policies obtruding technological obsolescence to IT resource mismanagement. In asset management institutions lacking oversight over network connectivity and peremptory policies lead to building excessive capacities--e.g. regarding long-haul fiber connections [12] and network access for particular end-users--as well as potentially inefficient procedures that obstruct the operations of companies. Consequently, internal factors sanctioning the adoption of specific AT solutions by traders yet need to be investigated.

## II. Trading Floor Architecture Overview

Trading floor architecture evolves continuously, both in terms of network solutions and applications in use. It is therefore necessary that upgrading a trading floor should involve minimal disruption to general system performance [13]. There are five groups of critical capacities and services to consider: (i) trading cluster connectivity and latency management, (ii) messaging and multicast, (iii) computing, orders matching, 'thin client' solutions, and trading resiliency, (iv) virtualization of data and application, and trading mobility, (v) data storage and access.

As shown in Figure 1, a trading environment typically has two components: a trading cluster with a ticker plant and algorithmic trading engines, and an end-user applications area. It also has to handle two types of traffic, both latency-sensitive. The first one is market data, unidirectional and typically delivered over a multicast from external feeds. The other one are trading orders, bidirectional, and measured in messages per second and Mbps.

The trading cluster and end-user applications areas can communicate via a message bus organized in topic streams, i.e. subsets of market data defined by criteria in such way that subscribers only receive the relevant information in order to facilitate their operations. The information subscribers base can be divided into topic groups mapped to one or multiple sub-topics, such as a ticker symbol, industry, or a certain class of financial instruments.

Fiber-optic communication is the typical method of transmitting information in OSMs. Notwithstanding, DMA implementations can also integrate other technologies. This is

P.Popławski is a Ph.D. student with the Institute of Telecommunications at the Warsaw University of Technology (email: ppoplaws @mion.elka.pw.edu.pl)

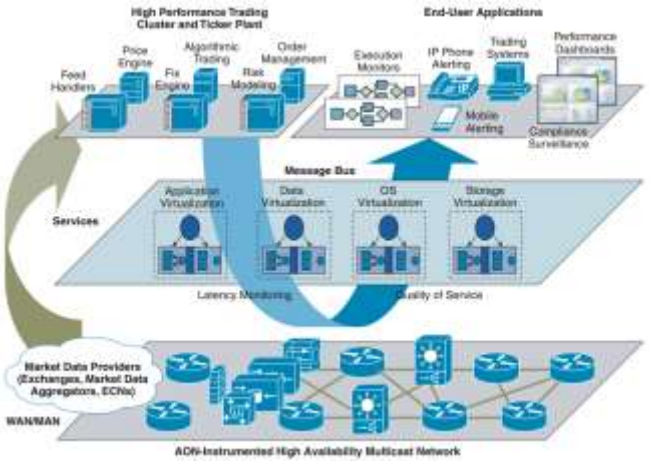


Fig. 1. Trading Floor Architecture. Source: [13]

particularly pertinent in the cases of large, institutional traders, whose simultaneous operations extended over several market centers optimize the potential for arbitration strategies. Notably, Marić [14] showed that in order to bridge large distances the solution that can best reduce transmission delays are microwave links, due to more direct routes and faster wave propagation.

The investment management industry tends to quickly adopt scalable standards, both for hardware and software. Hence speed inter-connect for the trading cluster, e.g. InfiniBand, 10G-PON data links or--in new installations--technologies for transmitting Ethernet frames at rates of up to 100 gigabits per second (40GbE and 100GbE) compliant with the IEEE 802.3ba-2010, 802.3bj-2014, or 802.3bm-2015 standards. High-speed messaging bus can operate with applications based on the Advanced Message Queuing Protocol and the solution for application acceleration without application re-code is remote direct memory access.

Network architecture is expected to proactively facilitate network management. Therefore it also faces several operational challenges such as correlating network events with application events for troubleshooting without adding latency to the trading traffic, or avoiding asymmetry that can be introduced by each node and link in a network and would affect the adequate time-phase accuracy, e.g. by using protocol-level full timing support [15]. Lastly, the latency monitoring system should synchronize with algorithmic trading engines to receive latency data and adapt to changing conditions, therefore optimizing the resilience of AT.

After latency management, another issue for the architecture of trading floor telecommunications network is the trade-off between scalability and required operational server capacity. Depending on the size and structure of a trading entity, a market data feed may need to reach hundreds or thousands of users. Therefore, solutions based on the Internet Protocol are preferred over the use of mere Transmission Control Protocol or User Datagram Protocol (UDP) broadcasts even when data is not supposed to leave a LAN and no need for internetworking appears. The needs for a separate socket and a sliding window in each recipient's server, as required by the TCP, or of individual copies of the information stream, as with the UDP, are thus avoided and both server and network capacities saved.

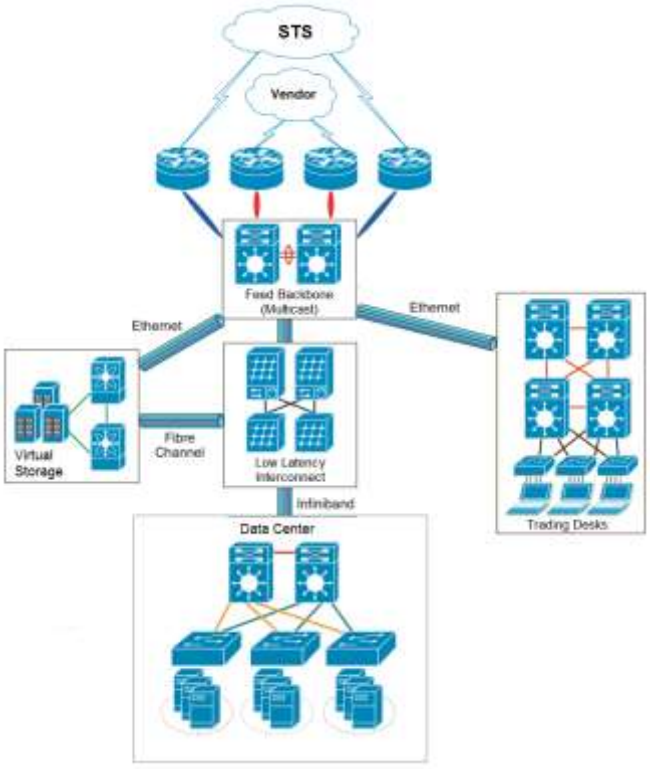


Fig. 2. Trading Environment

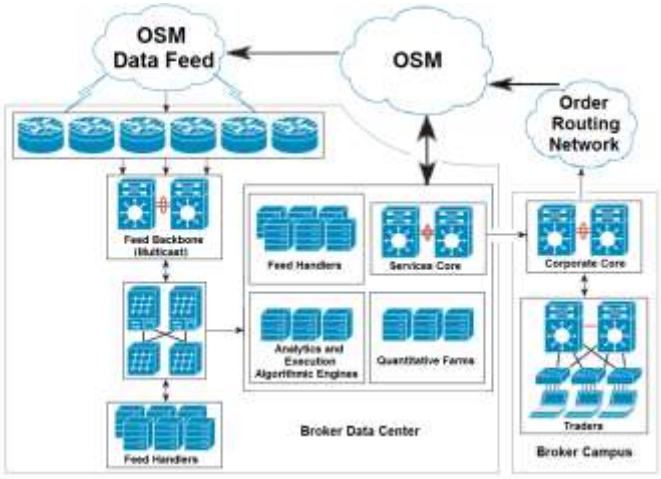


Fig. 3. AT Platform Model

The final problem concerning the architecture of a trading floor, especially pertinent for teams that are dispersed, located far from their trading cluster, serving various market centers simultaneously, or include numerous remote traders are the resilience, virtualization, and mobility of trading environment. To counteract any eventual disruptions or failures, e.g. related to modules, switches, or links, it may be built according to a campus-area network topology over dense wavelength division multiplexing or Dark Fiber, with non-stop forwarding and synchronous data replication.

Provided a sufficient computing capacity in the trading cluster, a widely accepted 'best practice' is to find the optimal feed normalization standard and order routing method, and use applications that treat end-users as 'thin clients' despite the impossibility to communicate with arbitrarily small

probability of error (Shannon reliability [16]) over channels that fail at random times [17]. IT support and resource management should thus focus on transmission quality, as the computing-related problems are soluble at the trading cluster.

With remote end-user applications area human traders are not able to take full advantage of DMA. Notwithstanding, the delays introduced by the human factor significantly exceed those related to technology. This implies the shift from a traditional human trader to computer-based AT system and from a client-facing broker to index and funds broker, consistent with market trends evolution.

## III. Direct Market Access

DMA is a range of solutions for electronic securities trading that enable traders to access the central order book of an OSM directly and in real time. For each trading operation, it allows to reduce transaction costs, time, and the likelihood of execution errors. Although DMA excludes any solution where access to an OSM requires an active presence of intermediary parties, network infrastructure can belong either to an independent provider, the OSM, or sell-side firms.

DMA solutions cover most of global financial markets. Instinet, the first passive computer-driven electronic communication network was created in 1969 and has been supporting DMA since 1980. However, it was not until the late 1990s that algorithmic trading of securities appended mainstream operations and until 2001 that the New York Stock Exchenge (NYSE) opened a fully automated securities exchange for trading stocks and options, ArcaEx. By the beginning of the following decade not only were DMA solutions present in global financial hubs but also in OSMs of emerging and peripheral markets. As of 2015, the latest major implementation of DMA was in the Johannesburg Stock Exchange with colocation services providing a roundtrip latency of 100-150 μs [18, 19].

Low latency, understood by traders as effective responsiveness to market events in a millisecond environment [20] and information security appear as the key drivers for the propagation of DMA. Furthermore, DMA enables the use of HFT, potentially precluding trading losses, as modeled by Hendershott and Riordan [2, 21] and Menkveld [22]; Virtu proved the deliverability of these theories in practice [23: 3].

There exist four basic models of DMA. 'Colocation,' where trading computers are located in the premises of an OSM and form a local area network (LAN) with the trading engines accountable for the central order book operations (central trading engines, CTEs) is the market connectivity solution typically associated with HFT. It provides the highest available speed and capacity. Aitken et al. [24] describe an endemic adoption of AT prior to its regularization within market centers and confirm that on most exchanges HFT predates the introduction of colocation services by at least eight months.

'Direct connection' between trading platforms and OSM servers as well as 'access via DMA provider' represent solutions based on metro area network (MAN) typologies. The DMA provider can either form a LAN with the CTEs of the OSM or be connected to the CTEs in a MAN.

## IV. Market Messages and Latency

In an AT scheme, there are two factors that can trigger activity. For orders programmed for execution or cancellation in predefined intervals it is the flow of time; otherwise it is a relevant market message. Market messages broadcast in the basic information feed of an OSM are the prices of securities and the volumes traded. Such extent of information usually suffices for technical analysis [25, 26, 27] and, consequently, for the execution of AT. Specific data structure and format depend on the OSM software provider and on data aggregation methodology. Certain economists [28, 29, 30, 31] relate the AT practice of canceling and resubmitting of orders to a systemic destruction of market liquidity. This suggests that not only does the quality of DMA represent an important determinant of a trader's operational efficiency but also has an impact on the markets on a macro level.

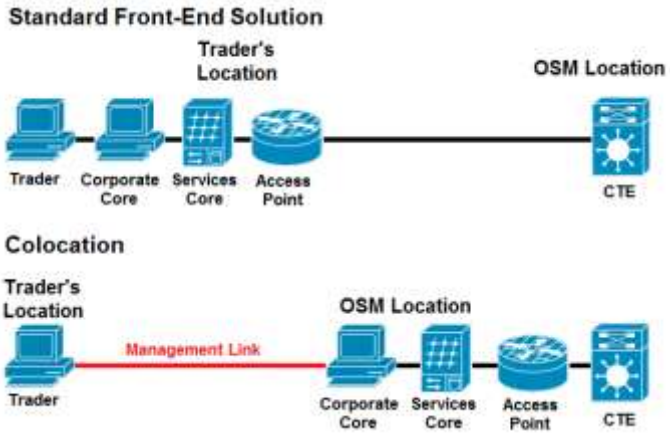


Fig. 4. A Comparison of a Standard Front-End Solution and Colocation

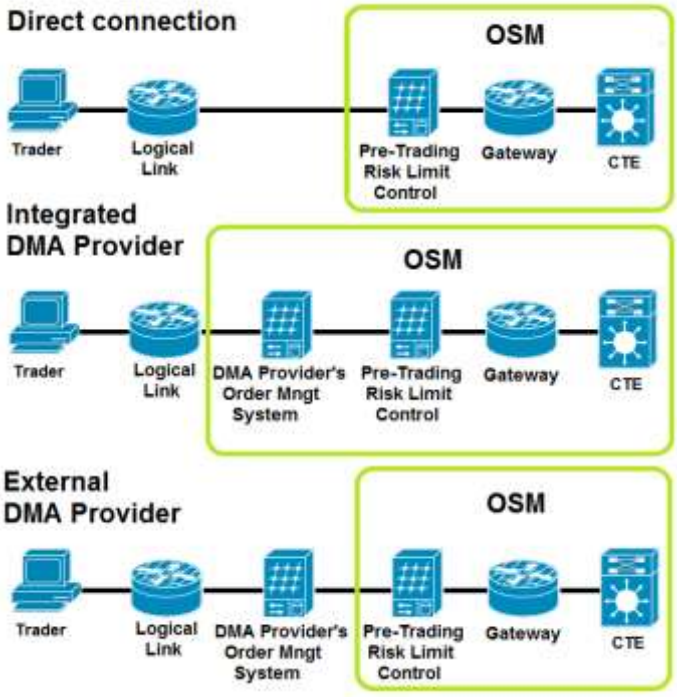


Fig. 5. DMA Connectivity Models

While liquidity accounts for the opportunity of executing a trade, latency is directly related to its profitability in an environment where prices change in time. Latency is also the only variable of operational efficiency objectively comparable between competing traders. Most of the accepted liquidity measures [32: 47-52] focus on the bid-ask spread values. Since discretionary price changes caused by particular trades tend to be lower in highly-liquid markets [33, 34] the cost of latency negatively correlates with market liquidity. The priority for a telecommunications network is therefore to ensure a possibly low time required to complete a transaction triggered by a market message.

Latency, or round-trip time (RTT), is a sum of: forward latency, matching engine latency, and outgoing latency. Freeman [35: 335] lists four factors that cause increases in network latency: (i) propagation delay, (ii) processing time and processing requirements, (iii) the number of message exchanges required to complete a transaction, and (iv) the bad quality of a circuit, e.g. when the circuit is noisy or many automatic repeat request exchanges occur. While the first two factors relate mostly to the quality of the trader's connection and hardware, the other two are equally influenced by its trading algorithms. The algorithms, in turn, depend on the trading strategy. In order to take advantage of low outgoing latency, the trading floor thus requires adequate computational and data access capacities. A basic element is database acceleration; this is achievable e.g. with field-programmable gate array (FPGA) based filters [36].

Moallemi and Sağlam [37] offer a theoretical formula for the cost of latency (CL) that can be applied for modeling the cost-efficiency of HFT solutions or the comparative efficiency of any AT solution against other solutions that apply equivalent trading strategies. Given these conditions, latency cost is a function of latency ($\Delta t$), price volatility ($\sigma$), and a bid-offer spread ($\delta$). The cost of latency can thus be interpreted as a cost of inefficiency comparable to unplanned obsolescence, with no impact on the operational costs of the trading party. The need to exert lower latency can be fully explained by changes of market conditions such as price fluctuations ($\sigma$) and an operational factor ($\delta$); the very latency ($\Delta t$) could be ignored unless higher than at direct competitors. The formula suggests that a market agent trading an asset should especially aim in minimizing the latency relevant for trading the asset if its price fluctuates significantly.

$$\underset{\Delta t \to 0}{C_L} = \frac{\sigma\sqrt{\Delta t}}{\delta}\sqrt{\log\frac{\delta^2}{2\pi\sigma^2\Delta t}} \quad (1)$$

In practice, the liquidity of an asset should also be recognized as a factor reducing the price change effect of each concluded transaction [38]. Including liquidity would therefore help to adjust Moallemi and Sağlam's formula to real-market conditions when modeling trading as a repetitive game, since trades under the same market conditions can have different values as a result of different quantities traded each time. Consequently, when several agents trade a particular asset, those with higher RTT would bear additional costs proportionately to the volume of trades executed. On a different note, the low frequency of trades executed on illiquid markets challenges the rationale for investing in RTT retrenchments since in such trading environment the risk of missing an opportune trade is reduced.

On top of in-house RTT optimization, competitive advantage over other traders may be gained by receiving a message from the market center before it reaches the competitors. Such phenomenon, known as latency arbitrage [39], is often considered as discriminatory against the traders who receive the message later. Usually, unless caused by technological factors, latency arbitrage is actively prosecuted by financial regulators.

An ultra low-latency environment can nonetheless corroborate certain otherwise illegal actions such as frontrunning, defined by Khan and Lu [40] as "trading by some parties in advance of large trades by other parties, in anticipation of profiting from the price movement that follows the large trade," without the trading party breaching the law. Network latency differences between competing traders or institutions offering brokerage, can enable a trading party to project the actions of other market participants prior to their execution by CTEs and thus to effectively circumvent existing regulations and pursue frontrunning.

## V. Automated Order Management and Market Externalities

Automated and human traders manage trading orders in two divergent ways. Automated trading engines submit an order and revisit it at fixed intervals with or without the occurrence of an event relative to a given security. Human traders usually respond to market events. The fixed intervals can be programmed either as a function of time, the number of registered trades of a security, its price fluctuations, or a combination of those variables; the accurate execution of such tactics by a human would not be viable. Furthermore, Hasbrouck and Saar [16] note that "an algorithm that repeatedly submits orders and cancels them within 10ms does not intend to interact with human traders (whose response time would probably take more than 200ms even if their attention is focused on this particular security)."

On the other hand, AT systems depend on reliable and uninterrupted connection with the order book of the OSM and on the lack of interference in communication with other data sources. DMA is therefore essential for AT. Without the visibility into which OSM offers the best conditions for price execution, trading engines may compound the increase of the volume of orders by issuing orders only to cancel and re-submit them to where the best price has been found. Such situation represents a challenge not only form the perspective of IT and telecommunications systems management but also for order management at the OSM centers as it increases the risk of market glitches and price shocks or crashes.

The competitive advantage of an AT trader over traditional traders, complemented by the large scale of AT, exerts negative externalities [41, 42] and results in the adverse selection of market agents despite incumbent measures taken to alleviate this phenomenon [43]. However, it is disputable whether human traders should compete with AT and HFT since in a market where both are present, each group can find a strategy that enables them to benefit form a different aspect of trading [44].

The technologies used for HFT produce a situation where a trading order can be canceled and re-issued several times prior to being officially registered. For RTT, as shown in Figure 6,

quotes on NYSE are registered by the Security Information Processor (SIP) after 250µs, while it only takes HFT engines colocated in the same data center in Mahwah, NJ 2µs to register them. Thus, more efficient communication systems for HFT also provide their users with the advantage of circumventing the delays in quotes transmission between securities exchanges. Having analyzed of over ten billion quotes and trades matched between direct feeds and the SIP with microsecond resolution timestamps in mid-2015, Nanex [45] suggests that HFT in Mahwah sees and can act on changes to quotes on Nasdaq--another stock exchange located in the New York City area--within 191µs, i.e. 125µs before the NYSE registers the change. Finally, while quotes on the electronic exchange BATS, located in Secaucus, NJ, are registered on NYSE's Tape A after 450µs, HFT engines colocated in the NYSE data center can react on them with the delay of only 125µs.

In automated markets latency and processing time contain valuable non-trade information about the price formation process in a trading system, as shown by Kirilenko and Lamacie [46]. After studying the Brazilian Securities, Commodities and Futures Exchange (Bovespa) they show that latency, random and highly variable--"an automated trading platform can take as little as 800 microseconds to process a traders message or as much as 80 milliseconds" and "variations are not well described by a bell-shaped distribution"--has a strong predictive power over both volatility and the volatility of volatility of a highly liquid asset over and above changes in message traffic. The speed and quality of communication inside a trading system can therefore be seen as a direct factor of market quality. One can assume Bovespa's colocation center opened in April 2014, over three years after the introduction of this DMA model to the OSM [47], and with no record of major technological or order management problems (as of October 2015) to be in line with global standards.

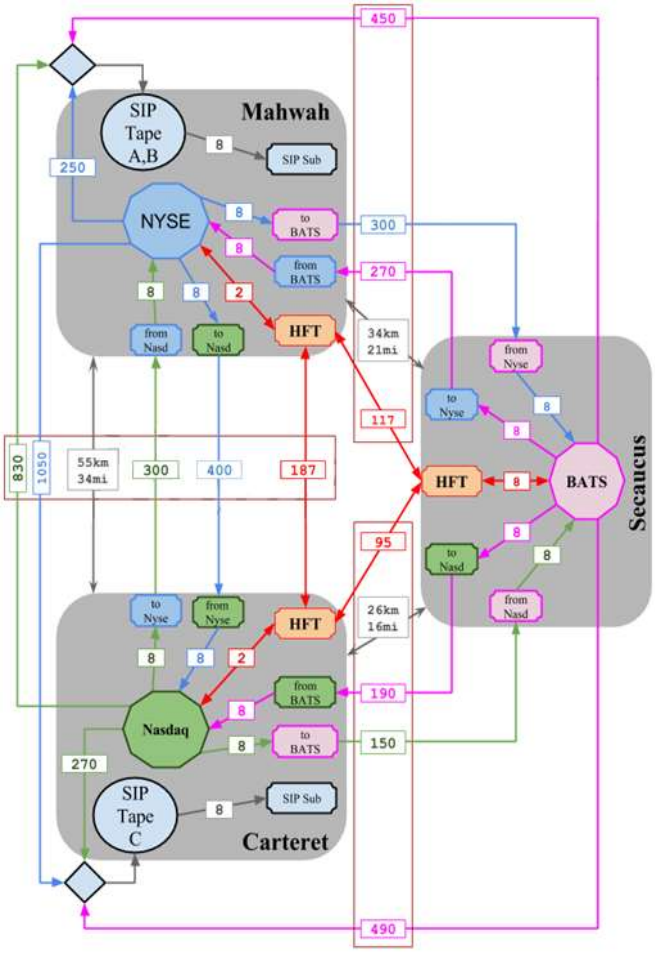


Fig. 6. Latencies in the Electronic Communication System of OSMs in Northern New Jersey, Measured in µs. Source: Nanex

## VI. Conclusions

The performance of an AT system depends predominantly on the responsiveness to market messages and on the efficiency of applied trading algorithms. A point of paramount importance for the operator of a trading floor is thus to minimize systemic latency levels and to manage the operational capacities in such ways that effective RTT remains possibly low. The optimization of the raw throughput and message rates for both market data and trading orders is a convergent objective. To achieve it, a trading floor needs adequate infrastructure and information management system.

Although particular traders can only perceive latency and the use of communications networks in an OSM as given external conditions, they should adjust their strategy to the trading conditions of their direct competitors. For market making purposes, trading systems use the rule of time priority. Accordingly, the order that comes first is first served and when two market agents use similar trading algorithms but different RTT, the reaction to a change in market price of the one with higher latency is delayed, trade execution may occur in less favorable conditions than its competitor's.

Agile management of trading systems and their OSM connections should provide a trader with an efficient use of resources. The possibility to separate brokers campuses from data centers, or to use alternative communication channels between OSMs only represent some of the areas where creative solutions can lead to gaining competitive advantage either by cost reduction or by excelling operationally.

When market price only reaches the level of a limit order accidentally, the execution of a trade commissioned via a higher-latency system is jeopardized. This emphasizes the importance of choosing the kind of DMA that best corresponds to a trader's strategy and external market conditions. While a specific DMA implementation can constraint the operational ability of a trader to outperform the competitors using more efficient DMA solutions, when trading on relatively illiquid markets the negative effects of higher latency may not prevail over the costs involved.

Finally, while the of human-based trading has lost its primary role and decoupled from AT, it should not be seen as the final stage of financial markets development. As long as artificial intelligence systems provide different solutions to specific problems than humans do, the advantage of the latter can abide. The perpetuation of several parallel trading technologies can lead to the formation of a new spectrum of market strategies.

## REFERENCES


[1] J.L. Teall, "Financial Trading and Investing," Waltham, MA: Elsevier Academic Press, 2013.

[2] T. Hendershott, R. Riordan, "High Frequency Trading and Price Discovery," working paper, University of California at Berkeley, 2011.

[3] J. Brogaard, "High frequency trading and its impact on market quality," working paper, SSRN, 2010.

[4] T. Hendershott, R. Riordan, "Algorithmic trading and information," Working Paper 09-08, NET Institute, 2010.

[5] A. Kirilenko, P. Kyle, M. Samadi, T. Tuzun, "The impact of high-frequency trading on an electronic market," working paper, University of Maryland, 2010.

[6] A. Admati, P. Peiderer, "The Value of Information in Speculative Trading," Research paper 782, Stanford University Graduate School of Business, 1984.

[7] S.J. Grossman, J.E. Stiglitz, "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, vol. 70, pp. 393-408, June 1984.

[8] A.W. Lo, "Hedge Funds: An Analytic Perspective," revised and expanded ed., Princeton, NJ: Princeton University Press, 2010.

[9] W.G. Lewellen, R.C. Lease, G.G. Schlarbaum, "Patterns of Investment Strategy and Behavior Among Individual Investors," *The Journal of Business*, vol. 50, no. 3, pp. 296-333, July 1977.

[10] B.M. Barber, T. Odean T., "The Internet and the Investor," *The Journal of Economic Perspectives*, vol. 15, no. 1, pp. 41-54, Winter 2001.

[11] G. Jackson, R. Deeg, "Comparing Capitalisms: Understanding Institutional Diversity and Its Implications for International Business," *Journal of International Business Studies*, vol. 39, pp. 540–561, 2008. DOI:10.1057/palgrave.jibs.8400375.

[12] J. Loveless, "Connecting to New Markets: Eight Costliest Network Mistakes," Traders Magazine Online News of September 18, 2015.

[13] A. Kessler, D. Malik, M. Risca, "Trading Floor Architecture," Cisco Systems Inc, 2008.

[14] I. Marić, "Low Latency Communications," presented at the Information Theory and Applications Workshop (ITA 2013), San Diego, CA, February 10-15, 2013, arXiv:1302.5662 [cs.IT].

[15] S. Siu, W.H. Tseng, H.F. Hu, Sh.Y. Lin, Ch.Sh. Liao, Y.L. Lai, "In-Band Asymmetry Compensation for Accurate Time/Phase Transport over Optical Transport Network," *The Scientific World Journal*, vol. 2014, Article ID 408613. DOI: 10.1155/2014/408613.

[16] S.W. Ho, "On the interplay between Shannon's information measures and reliability criteria," 2009 IEEE International Symposium on Information Theory - ISIT, June 28-July 3 2009. DOI: 10.1109/ISIT.2009.5205836.

[17] L.R. Varshney, S.K. Mitter, V.K. Goyal, "An Information-Theoretic Characterization of Channels That Die," *IEEE Transactions on Information Theory*, vol. 58, no. 9, pp. 5711-5724, 2012.

[18] JSE, "The lowest-latency connection to JSE markets," Johannesburg: The Johannesburg Stock Exchange, April 2014.

[19] JSE, "JSE launches high tech Colocation Centre," Johannesburg: The Johannesburg Stock Exchange, May 14, 2014.

[20] J. Hasbrouck, G. Saar, "Low-Latency Trading," *Journal of Financial Markets*, vol. 16, no. 4, pp. 646–679, November 2013.

[21] J. Brogaard, T. Hendershott, R. Riordan, "High Frequency Trading and Price Discover," The European Central Bank Working Paper Series, No. 1602 / November 2013.

[22] A.J. Menkveld "High frequency trading and the new market makers," *Journal of Financial Markets*, vol. 16, no. 4, pp. 712–740, November 2013.

[23] Virtu, Form S-1: Registration Statement under the Securities Act of 1933 filed by Virtu Financial Inc. on March 10, 2014.

[24] M. Aitken, D. Cumming, F. Zhan, "Trade size, high-frequency trading, and colocation around the world," *The European Journal of Finance*, no. 12/2014. DOI:10.1080/1351847X.2014.917119.

[25] A.W. Lo, H. Mamaysky, J. Wang "Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation," *The Journal of Finance*, vol. 55, no. 4, pp. 1705-1765, August 2000.

[26] L. Blume, D. Easley, M. O'Hara, "Market Statistics and Technical Analysis: The Role of Volume," *The Journal of Finance*, vol. 49, no. 1, pp. 153-181, March 1994.

[27] R.D. Edwards, J. Magee, W.H.C. Bassetti, "Technical Analysis of Stock Trends," ninth edition, CRC Press, 2007.

[28] Th. Foucault, O. Kadan, E. Kandel, "Limit Order Book as a Market for Liquidity," *The Review of Financial Studies*, vol. 18, no. 4, pp. 1171-1217, Winter 2005.

[29] C.M. Jones, "What Do We Know About High-Frequency Trading?," Columbia Business School Research Paper No. 13-11, 2013.

[30] M. O'Hara, "What Is a Quote?," *The Journal of Trading*, vol. 5, no. 2, pp. 10-16, Spring 2010. DOI: 10.3905/JOT.2010.5.2.010.

[31] I. Poirier, "High-frequency trading and the flash crash: structural weaknesses in the securities markets and proposed regulatory responses," Hastings Bus. LJ 445, 2012.

[32] R.J. Riordan, "The Economics of Algorithmic Trading," doctoral thesis presented in the Universität Karlsruhe (TH), 2009.

[33] T. Chordia, R. Roll, A. Subrahmanyam, "Market Liquidity and Trading Activity," *The Journal of Finance*, vol. 56, no. 2, pp. 501-530, April, 2001.

[34] M.K. Brunnermeier, L.H. Pedersen, "Market Liquidity and Funding Liquidity," NYU Stern Working Paper Series, SC-AM-05-06, 2005.

[35] R. Freeman, "Fundamentals of Telecommunications," John Wiley & Sons, 1999.

[36] J. Teubner, L. Woods, Ch. Nie, "XLynx—An FPGA-based XML Filter for Hybrid XQuery Processing," *ACM Transactions on Database Systems*, vol. 38, no. 4, Article XX, 2013. DOI: 10.1145.

[37] C.C. Moallemi, M. Sağlan, "The Cost of Latency in High-Frequency Trading," *Operations Research*, vol. 61, no. 5, pp. 1070–1086, 2013.

[38] Y. Amihud, H. Mendelson, "Liquidity and Asset Prices: Financial Management Implications," *Financial Management*, vol. 17, no. 1., pp. 5-15, Spring, 1988.

[39] J.F. Egginton, B.F. Van Ness, R.A. Vann Ness, "Quote Stuffing," 2014. Available at SSRN: DOI: 10.2139/ssrn.1958281.

[40] M. Khan, H. Lu, "Do Short Sellers Front-Run Insider Sales?," *The Accounting Review*, vol. 88, no. 5, pp. 1743-1768, September 2013.

[41] B. Bias, Th. Foucault, S. Moinas, "Equilibrium High Frequency Trading," Proceedings from the fifth annual Paul Woolley Centre conference, London School of Economics, 2011.

[42] B. Jovanovic, A. Menkveld, "Middlemen in limit order markets," working paper, New York University, 2010.

[43] Th. Foucault, A. Roell, P. Sandas, "Market making with costly monitoring: an analysis of the SOES controversy," *Review of Financial Studies*, no. 16, pp. 345-384, 2003.

[44] Á. Cartea, J. Penalva, "Where is the Value in High Frequency Trading?," 2011. DOI:10.2139/ssrn.1712765.

[45] Hunsader E.S.--Nanex LLP twitter.com/nanexllc/status/632610118874501120, August 15, 2015. Retrieved on October 9, 2015.

[46] A. Kirilenko, G. Lamacie, "Latency and Asset Prices," Working Paper, 2015.

[47] BM&F Bovespa, Ofício circular 001/2011-DP, São Paulo: Brazilian Securities, Commodities and Futures Exchange, 2011.